DOCUMENT RESUME

ED 349 865 HE 025 782

AUTHOR Pike, Gary R.
TITLE The Relationship between Perceived Learning and Satisfaction with College: An Alternative View. AIR 1992 Annual Forum Paper.
PUB DATE May 92
NOTE 27p.; Paper presented at the Annual Forum of the Association for Institutional Research (32nd, Atlanta, GA, May 10-13, 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Paper (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Alumni; College Seniors; Graduate Surveys; Higher Education; Institutional Research; *Learning; *Models; *Participant Satisfaction; *Student Attitudes; *Student Experience; Surveys
IDENTIFIERS *AIR Forum; *Halo Effect; University of Tennessee Knoxville

ABSTRACT

This study evaluated two competing explanations of the learning-satisfaction relationship in the college experience. The first model represents a true relationship between reported learning and satisfaction, while the second model treats this relationship as an artifact of a halo effect. The study took place at the University of Tennessee, Knoxville (UTK) and involved data collected in 1988 using a campus-wide assessment program where all seniors were administered a survey designed to elicit information about perceptions of their college experiences, satisfaction with college, and learning and development during college. In 1990, the same subjects were mailed an alumni survey containing many of the items from the senior survey, with 989 responses. The data were evaluated according to goodness-of-fit of the models, invariance of the measurement models across senior and alumni data, and multitrait-multimethod tests of the relationships between latent variables representing the same constructs for the senior and alumni data. Analyses revealed that treating the learning-satisfaction relationship as an artifact of a halo effect provided the best representation of the data. Although not conclusive, results suggested that educational researchers and assessment practitioners should be careful in interpreting self-reports of learning and development, particularly as they relate to satisfaction with college. Included are 47 references. (JB)

#21

THE RELATIONSHIP BETWEEN PERCEIVED LEARNING AND SATISFACTION WITH COLLEGE: AN ALTERNATIVE VIEW

Gary R. Pike

Associate Director

Center for Assessment Research and Development

University of Tennessee, Knoxville

1819 Andy Holt Avenue

Knoxville, TN 37996-4350

(615) 974-2350

Running Head: LEARNING AND SATISFACTION

Paper presented at the annual forum of the Association for Institutional Research, Atlanta, Georgia, May 10-13, 1992

for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Second Annual Forum of the Association for Institutional Research held at the Atlanta Hilton & Towers, Atlanta, Georgia, May 10-13, 1992. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Jean Endo
Chair and Editor
Forum Publications
Editorial Advisory Committee

#27

Abstract

Student and alumni reports of learning and development during college play an important role in research on educational outcomes. An intriguing finding of this research is the positive relationship between perceived learning and satisfaction with college. While studies have documented an association between perceptions of learning and satisfaction, the nature of the relationship is not clearly defined. This study evaluates two competing models of perceived learning and satisfaction. The first model represents a true relationship between learning and satisfaction, while the second treats the relationship as an artifact of a halo effect. Data came from subjects who completed learning and satisfaction questions as seniors and again two years after graduation. Analyses revealed that treating the learning-satisfaction relationship as an artifact of a halo effect provided the best representation of the data. Although not conclusive, results suggested that educational researchers and assessment practitioners should be careful in interpreting self-reports of learning and development, particularly as they relate to satisfaction with college.

THE RELATIONSHIP BETWEEN PERCEIVED LEARNING AND SATISFACTION WITH COLLEGE: AN ALTERNATIVE MODEL

Student and alumni reports of learning and development during college play an important role in research on educational outcomes. Feldman and Newcomb (1969), Lenning (1977), and Pascarella and Terenzini (1991) devoted considerable attention to how students' college experiences affect their perceptions of learning and development. Perceptions of learning and development also are an important component of many assessment programs. The ACE/Winthrop (1990) survey found that 65% of colleges and universities relied on measures of learning and development to assess their education programs, many of them using surveys of enrolled students or alumni.

Research has consistently found that dimensions of learning and development are positively related to each other and to satisfaction with college (Pace, 1987, 1988; Pike, 1990a, 1991, 1992; Terenzini, Pascarella, and Lorang, 1982). The learning-satisfaction relationship is particularly significant because satisfaction has been linked to persistence and academic performance (Aitken, 1982; Bean, 1980; Bean and Bradley, 1986; Munro, 1981; Pascarella and Chapman, 1983; Pascarella, Smart, and Ethington, 1986; Pike, 1989, 1990b; Terenzini and Pascarella, 1977). While research has documented an association between perceptions of learning and satisfaction, the nature of this relationship is not clearly defined. This study evaluates two competing explanations of the learning-satisfaction relationship. The first model represents a true relationship between reported learning and satisfaction, while the second model treats this relationship as an artifact of a halo effect.

Background

Arguing that the outcomes of a college education are multifaceted is hardly controversial. Several authors have proposed multidimensional typologies of college outcomes (Alexander and Stark, 1986; Astin, 1974; Bowen, 1977; Ewell, 1984), and the most recent review of research on how college affects students identified as many as ten dimensions of college outcomes (Pascarella and Terenzini, 1991). In an early study, Pascarella and Terenzini (1978) examined the responses of 528 second semester freshmen to eleven items concerning gains made during college. Factor analysis identified two dimensions: personal development and intellectual development. Alpha reliability coefficients for the two

scales were .80 and .74 respectively. Subsequent research by Pascarella, Terenzini, and their colleagues, using an expanded list of gains made during college, identified additional dimensions of learning and development. Factor analysis of the responses of 469 second semester freshmen to a list of eighteen questions identified four dimensions of learning and development: personal growth, academic process, academic content, and future preparation (Terenzini, Pascarella, and Lorang, 1982). In this study, the researchers reported that the correlations among factor scores ranged from .37 to .58.

One of the most widely used instruments for measuring students' perceptions of gains made during college is the College Student Experiences Questionnaire (CSEQ). In the technical manual for the instrument, Pace (1987) reported results for more than 25,000 undergraduates from 74 institutions who completed the CSEQ between 1983 and 1986. Factor analysis of responses to the learning and development items in the survey identified five dimensions: personal and social development; science and technology; general education, literature, and arts; intellectual skills; and vocational preparation. Correlations between the dimensions were low to moderate, ranging from .18 to .52, with the intellectual skills dimension being most highly correlated with the other learning and development dimensions.

Surveys of alumni perceptions of learning and development during college also are an important source of data concerning college outcomes. Valiga (1982) analyzed the responses of almost 13,000 former students from 55 colleges and universities who completed the ACT Alumni Survey between 1980 and 1982. He identified six dimensions of learning and development: learning skills; science, mathematics, and problem solving skills; interpersonal and group dynamics skills; humanistic outcomes; basic communication skills; and life skills. In a replication of Valiga's work, Graham and Cockriel (1989) examined the responses of more than 50,000 alumni from 172 institutions to the ACT Alumni survey. They also identified six dimensions of learning and development: planning and organization skills, analytical thinking skills, self-directed learning skills, humanistic or artistic appreciation skills, communication skills, and consumer awareness and responsibility skills.

In Tennessee, funding supplements to public postsecondary institutions' budgets for instruction are partly determined by alumni responses to a series of learning and development questions. Pike (1990) examined data for almost 2500 alumni from five institutions who responded to the Tennessee

alumni survey in 1988. Restricted (confirmatory) factor analysis of 15 items revealed the existence of four dimensions of learning and development: personal and social development, growth in quantitative skills, growth in communication skills, and growth in cultural understanding. In a related study, Pike (1991) sought to validate the use of all 21 learning and development questions from the Tennessee survey as a measure of program quality at the University of Tennessee, Knoxville. Using data from two randomly-selected samples of alumni surveyed in 1988 and 1990, he identified four dimensions: development of verbal skills, development of quantitative skills, gains in knowledge outcomes (the arts and cultural understanding), and attitudes and judgements (personal development).

In a 1992 report, Pike examined the stability of students' perceptions of their learning and development at graduation and two years later. Using data from 916 graduates from the University of Tennessee Knoxville, he found that a four-factor structure provided a satisfactory explanation of the data for seniors and alumni, and that the factor structure was stable over time. Pike also found that the four factors representing perceived learning and development were moderately intercorrelated. Correlations ranged from .50 to .73, with the highest correlations being found between personal development and the other factors. He concluded that a higher-order construct was operating across all four dimensions of learning and development.

Research also has shown that perceptions of learning and development are positively related to satisfaction. In a report summarizing national results for the CSEQ, Pace (1988) found that all five dimensions of reported learning and development were positively related to satisfaction. Pike (1990a) found that the dimensions of personal and social development and growth in math and science from the Tennessee Alumni Survey were positively related to academic satisfaction. In addition, he found that personal and social development, as well as growth in cultural understanding, were positively related to social satisfaction. Personal and social development and growth in science and math were positively related to satisfaction with employment.

In addition to these studies, research has provided indirect evidence of a relationship between satisfaction and perceived learning. Terenzini, Pascarella, and Lorang (1982) examined the relationship between academic and social involvement and satisfaction. They found that both academic and social

involvement were positively related to all four dimensions of growth and development in their survey. What is significant about these results is that the involvement scales included several questions related to satisfaction with social and classroom experiences. Similar results have been found in other studies by these researchers (Terenzini, Theophilides, and Lorang, 1984a, 1984b; Terenzini and Wright, 1987a, 1987b)

Although the relationship between perceived learning and satisfaction is well established, the precise nature of this relationship remains unclear. Pike assumed that the direction of the relationship was such that perceptions of learning and development influenced satisfaction. In contrast, the research by Pascarella, Terenzini, and their colleagues assumed that satisfaction, as a component of academic and social involvement, influenced perceived learning and development. Pace (1988) noted that the direction of the learning-satisfaction relationship was impossible to determine. This conclusion has been supported by research on the relationship between satisfaction and college grades which found that there was a reciprocal relationship between satisfaction and grades (Bean and Bradley, 1986; Pike, 1989b, 1990).

Irrespective of the direction of the learning-satisfaction relationship, all of the studies cited in this review assumed that the relationship was "true." That is, correlations among satisfaction and learning dimensions were the product of a meaningful higher-order construct. An alternative explanation is that the observed correlations between satisfaction and perceptions of learning and development were the product of response consistencies across survey items (Bradburn, 1983). For example, students who were more satisfied may have reported that they learned more than students who were less satisfied. According to this scenario, the observed correlations among satisfaction and dimensions of perceived learning are a product of a constant error or "halo" effect (Thorndike, 1920). The remainder of this paper examines these two models in light of senior and alumni responses to the Tennessee Alumni Survey.

Theoretical Models

Figure 1 presents a model representing the assumption that the observed relationship between perceived learning and satisfaction is true. Observed responses to satisfaction and learning items (Y_1 ...

Y_9) are the product of three factors (η_1 ... η_3) and measurement errors (ε_1 ... ε_9). Relationships among satisfaction (η_1) and perceived learning (η_2 and η_3) are represented by a second-order factor (ξ_1) and the paths between the first and second-order factors ($\gamma_{1,1}$ ... $\gamma_{3,1}$) indicate the strength of the relationships. In the model, those aspects of satisfaction and perceived learning that are unrelated (unique) are represented by zeta (ζ_1 ... ζ_3). In order for this model to be identified, the paths from the first-order factors to the measures variables, along with the variance for the second-order factor, must be fixed at unity (Jöreskog and Sörbom, 1989).

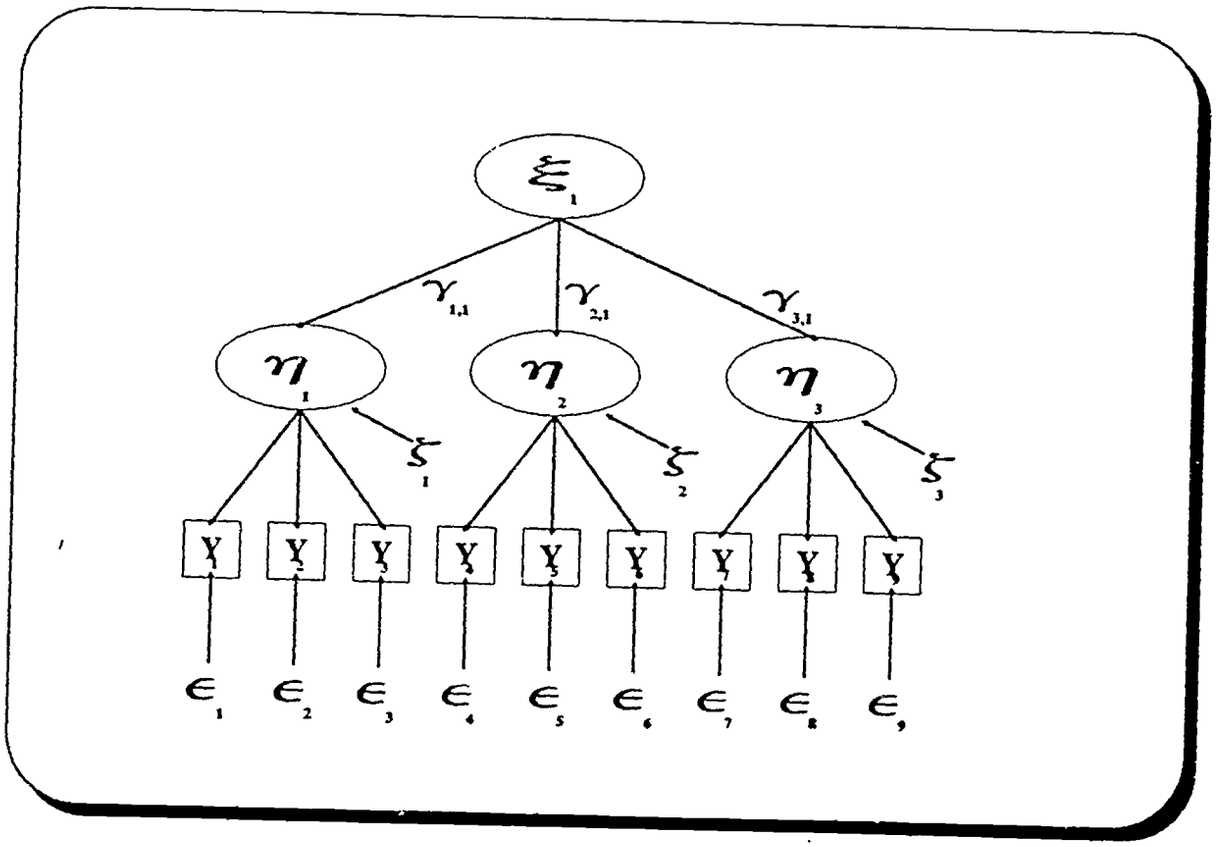

Figure 1:

A Hypothetical Model Representing the Relationship Between Satisfaction and Perceived Learning as "True"

Figure 2 depicts the alternative model of the relationship between satisfaction and perceived learning. In the model, observed variables (Y_1 ... Y_9) are assumed to be the product of a common factor

(η_1), representing the halo effect, three unique factors representing satisfaction (η_2) and perceived learning (η_3 and η_4), and measurement errors (ε_1 ... ε_9). Unlike the first model, in which the common factor represents relationships between the latent variables (factors), the common factor in the alternative model represents the intercorrelations among observed variables. Thus, correlations are attributable to response consistencies across the observed variables, rather than a true relationship between learning and satisfaction (the latent variables). Due to constraints that must be imposed on this model in order for it to be identified, paths from the three unique factors to the observed variables, and the variance for the common factor, must be fixed to unity. In addition, the common and unique factors must be uncorrelated (Gold and Muthén, 1991).

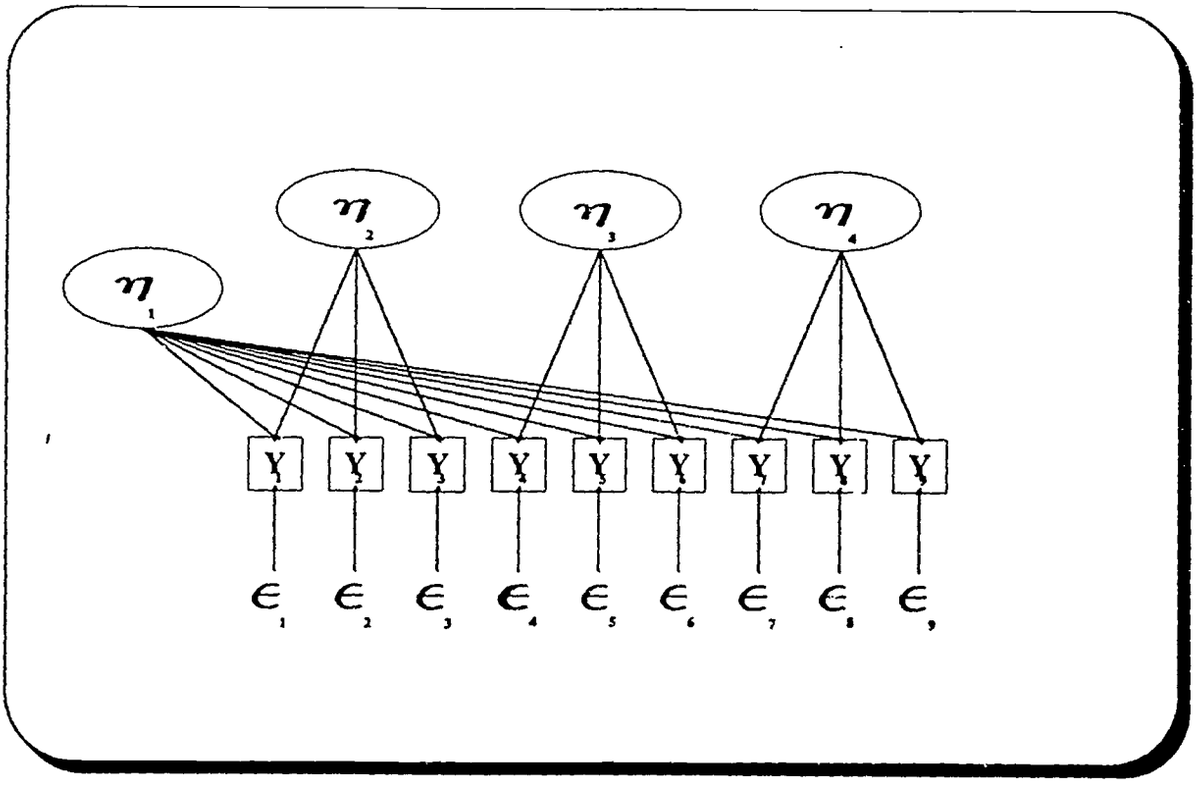

Figure 2:

A Hypothetical Model Representing the Relationship Between Satisfaction and Perceived Learning as a Halo Effect

Methods

Subjects

The setting for this research was the University of Tennessee, Knoxville (UTK), the state's public research university. In 1988, as part of the campus-wide assessment program, all seniors completing the general education testing requirement were administered a survey designed to elicit information about perceptions of their college experiences, satisfaction with college, and learning and development during college. In 1990, the same subjects were mailed an alumni survey containing many of the items from the senior survey. The response rate for the senior survey was essentially 100% (foreign students are excluded from the testing requirement), while the response rate for the alumni survey was slightly greater than 50%.

Complete data were obtained for 989 individuals. Approximately 50% of the subjects were males and 96% were white. In 1988, the mean age for the group was 23.6, and 94% of the subjects reported attending UTK full time. The mean entering Enhanced ACT Assessment composite score for the group was 22.3, and the mean cumulative grade point average during college was 2.94. Analysis of variance results indicated that the subjects included in the study were not substantively different from the total population of 1988 graduates in terms of their background characteristics (age, gender, and race), enrollment status, entering ability levels, or academic achievement during college.

Instrument

Data for this study were derived from responses to fifteen questions, common to both the senior and alumni surveys. Previous research found that these questions represent five dimensions: satisfaction (three questions), verbal skills (three questions), quantitative skills (three questions), personal development (four questions), and the arts and humanities (two questions) (Pike, 1990, 1991, 1992). In both surveys, subjects were asked to indicate their satisfaction with their academic, social, and overall experiences. Response options were "very dissatisfied," "dissatisfied," "satisfied," and "very satisfied." Because of substantial negative skewness in these responses, the "very dissatisfied" and "dissatisfied" responses were combined. In responding to the learning and development items, subjects were asked to indicate the degree to which their education at UTK added to their skills in the areas represented by the

twelve learning and development questions. Three response options were provided: "very little," "somewhat," and "very much."

Data Analysis

Three criteria were used to evaluate the competing models of senior and alumni responses to the satisfaction and perceived learning questions: (1) goodness-of-fit of the models to senior and alumni data; (2) invariance of the measurement models across senior and alumni data; and (3) multitrait-multimethod tests of the relationships between latent variables representing the same constructs for the senior and alumni data.

The warrant for the first criterion is obvious. The superior model is the one which provides the best explanation of the observed data. The basis for the second criterion may be less obvious. Nevertheless, it is a critical element in longitudinal students of learning and development (Pike, 1992). Unless the invariance of the outcomes measures can be established, it is impossible to determine whether differences over time represent real changes or whether they are products of different constructs being measured at different points in time (Byrne, Shavelson, and Muthén, 1989). Both Cunningham (1991) and Horn (1991) noted that evidence of measurement invariance requires, at a minimum, that factor loadings and uniquenesses in the measurement model be invariant. Pike (1991, 1992) observed that requiring that the factor loadings, uniquenesses, and factor variance-covariance (or correlation) matrices to be the same represents the most stringent test of measurement invariance.

The third criterion, multitrait-multimethod evaluation of the relationship between latent variables at two points in time, provides a basic test of validity that has been used extensively in evaluating outcomes measures (Messick, 1989; Pike, in press). This criterion requires that latent variables be more strongly related to their counterparts at different points in time than they are to other latent variables. Due to requirements for identified models, a more stringent criterion was imposed in this research. Specifically, identical factors in the senior and alumni data were assumed to be significantly correlated, while different factors were assumed to be uncorrelated.

Restricted (confirmatory) factor analysis procedures were used to test the true and halo effect models. Using item responses, rather than scale scores, in a restricted factor analysis can create serious

problems when responses are not normally distributed (Bernstein and Teng, 1989). Previous research using these data also revealed that the learning and development items were ordinal and that item responses contained substantial negative skewness (Pike, 1991, 1992). Using the PRELIS computer program, it is possible to infer underlying normally-distributed interval-level variables (Y*) from the ordinal variables (Y). Because the scale of the underlying interval-level variables is arbitrary, the correlation matrix should be analyzed instead of the variance-covariance matrix (Jöreskog and Sörbom, 1989). In the present research, matrices of polychoric correlations were calculated and then analyzed using weighted least squares techniques (Jöreskog and Sörbom, 1989). This approach allowed interval-level variables to be inferred from the ordinal measures, controlled for restriction of range in the correlation coefficients, and adjusted for skewness in the observed measures.

For each criterion, several models were specified and evaluated for goodness-of-fit. In order to evaluate the fit of the models to the observed data, a null model was first specified and tested. This model provided a baseline against which other models could be compared. The null model contained five latent variables identical to those in Figures 1 and 2. However, in the null model the latent variables were uncorrelated. This model differs markedly from the traditional null model of uncorrelated measured variables. It was selected because the measurement structure for the observed variables had been well established and because the goal was to test different models of the relationships and variables.

The second and third models evaluated were identical to the true and halo effect models in Figures 1 and 2. For the first criterion, analyses involved evaluating the improvement in goodness-of-fit achieved by moving from the null model to each of the alternative models. Tests of the second criterion utilized the same three models, except that factor loadings, uniquenesses, and factor variances, covariances, and correlations were assumed to be invariant across the senior and alumni data. Tests of the third criterion also utilized the true and halo effect in Figures 1 and 2. For these analyses, correlations between the same latent variables representing senior and alumni responses were free to vary. Correlations among different latent variables were all fixed at zero.

Three indices of improvement in goodness-of-fit were utilized: (1) the traditional chi-squared statistic; (2) the Relative Noncentrality Index (McDonald and Marsh, 1990); and (3) the Tucker Lewis

Index (Marsh, Balla, and McDonald, 1988). The chi-squared statistic represents the traditional measure of model goodness-of-fit; however, it is strongly influenced by sample size (i.e., small samples tend to produce artificially low chi-squared values while large samples produce inflated chi-squared values, indicating poor fitting models) (Marsh, Balla, and McDonald, 1988). The Relative Noncentrality Index (RNI) and the Tucker-Lewis Index (TLI) avoid problems created by sample size, with the RNI providing an unbiased estimated similar to the Bentler-Bonett Index and the TLI providing a measures accounting for model parsimony. Both the RNI and TLI produce values ranging from 0 to 1, with a value of approximately .9 representing a good fitting model. Because the traditional null model of uncorrelated measured variables was not used, the .9 value recommend for the RNI and TLI may be relaxed.

Results

Table 1 presents the goodness-of-fit results for the three models of the senior data. As can be seen from the information in the table, the null model provided a very poor representation of the observed data (χ^2 = 4,840.43; df = 100; $p < .001$). While both the true model and the halo effect model also produced significant chi-squared values (χ^2 = 461.06; df = 95; $p < .001$ and χ^2 = 327.86; df = 85; $p < .001$ respectively), this represented a substantial improvement over the null model. Both the RNI and TLI values for the halo effect model (.93 and .94) were somewhat better than the corresponding values for the true model (.90 and .92).

Table 1:

Goodness-of-Fit Results for the Senior Data

Model	df	χ^2	RNI	TLI
Null model	100	4,840.43*	---	---
True model	95	461.06*	0.90	0.92
Halo effect model	85	327.86*	0.93	0.94

* $p < .001$

Similar results were found for models of the alumni data. These results are presented in Table 2. Again, the null model provided a very poor representation of the observed data (χ^2 = 5,009.17; df = 100; $p < .001$). Although neither the true nor the halo effect models provided as good a representation of the alumni data as they did the senior data (χ^2 = 816.02; df = 95; $p < .001$ and χ^2 = 479.29; df = 85; $p < .001$ respectively), they did result in substantial improvements in chi-squared goodness-of-fit. The RNI and TLI for the halo effect model (.89 and 89) were very close to the recommended value of .90, and they were substantially better than the corresponding values for the true model (.84 and .84).

Table 2:

Goodness-of-Fit Results for the Alumni Data

Model	df	χ^2	RNI	TLI
Null model	100	5009.17*	---	---
True model	95	816.02*	0.84	0.84
Halo effect model	85	479.29*	0.89	0.89

* $p < .001$

Tests of model invariance produced results consistent with those found for the senior and alumni data independently. These results are presented in Table 3. As expected, a null model representing uncorrelated latent variables provided a very poor representation of the observed data (χ^2 = 9883.07; df = 220; $p < .001$). While both the true and halo effect models also produced statistically significant chi-squared values (χ^2 = 1448.79; df = 215; $p < .001$ and χ^2 = 1089.93; df = 205; $p < .001$), these values represented substantial improvements over the null model. The RNI and TLI coefficients were substantially better for the halo effect model (.89 and .90) than for the true model (.85 and .87), and they were generally consistent with the values recommended for an acceptable model

Table 3:

Goodness-of-Fit Results for Tests of Model Invariance

Model	df	χ^2	RNI	TLI
Null model	220	9883.07*	---	---
True model	215	1448.79*	0.85	0.87
Halo effect model	205	1089.93*	0.89	0.90

* $p < .001$

Table 4 presents the results for the goodness-of-fit results for the multitrait-multimethod analyses. Results indicated that the halo effect model provided the best representation of the observed data. Chi-squared results indicated the null model provided a very poor representation of the observed data (χ^2 = 11,339.51; df = 425; $p < .001$), as did the true model (χ^2 = 2116.68; df = 409; $p < .001$) and the halo effect model (χ^2 = 1700.70; df = 389; $p < .001$). However, the true and halo effects models were vastly superior to the null model. As with the other tests, the RNI and TLI coefficients for the halo effect model (.85 and .87) were superior to those for the true model (.81 and .84.

Table 4:

Goodness-of-Fit Results for the Multitrait-Multimethod Analyses

Model	df	χ^2	RNI	TLI
Null model	425	11,339.51*	---	---
True model	409	2116,68*	0.81	0.84
Halo effect model	389.00	1700.70*	0.85	0.87

* $p < .001$

While the goodness-of-fit results provided consistent evidence of the superiority of the halo effect model, examination of the variance in observed variables explained by the two competing models suggested that goodness-of-fit results should be interpreted with caution. The variance in observed

variables explained by the competing models is contrasted in Table 5. Data in the table were taken from the models used to test measurement invariance. Because both common and unique factors were used to explain the relationships among observed variables in the halo effect model, Table 5 contains estimates of explained variance for both types of factors.

Table 5:

Common and Unique Variance Estimates for the Observed Measures of Satisfaction and Learning

	True	Halo		
Measure	Total	Common	Unique	Total
Academic satisfaction	0.62	0.14	0.42	0.56
Social satisfaction	0.62	0.17	0.42	0.59
Overall satisfaction	0.62	0.29	0.42	0.71
Writing effectively	0.63	0.36	0.23	0.59
Speaking effectively	0.63	0.37	0.23	0.60
Understanding written information	0.63	0.51	0.23	0.74
Defining and solving problems	0.54	0.49	0.24	0.73
Understanding mathematical concepts	0.54	0.19	0.24	0.43
Understanding graphic information	0.54	0.31	0.24	0.55
Learning on your own	0.54	0.46	0.04	0.50
Working cooperatively in a group	0.54	0.49	0.04	0.53
Planning and carrying out projects	0.54	0.58	0.04	0.62
Ability to lead or guide others	0.54	0.49	0.04	0.53
Understanding different philosophies/cultures	0.56	0.18	0.38	0.56
Understanding and appreciating the arts	0.56	0.18	0.38	0.56

Examination of the total variance in the observed variables explained by the true and halo effect models revealed little substantive difference in the models' explanatory power. The true model had greater explanatory power for eight observed variables, while the halo effect model had greater explanatory power for five variables. The explanatory power of the models was identical for the two items measuring learning in the arts and humanities. On average, the halo effect model explained 1% more of the variance in each of the observed variables than did the true model.

Decomposition of the total variance into common and unique components produced interesting results. Counter to expectations, the common factor was able to explain less variance in the observed satisfaction measures than was the unique satisfaction factor. The unique factor representing learning in the arts and humanities also explained a majority of the variance in the observed arts and humanities measures. The unique factor labeled development of quantitative skills was responsible for a majority of the explained variance in the item concerning mathematics skills, indicating that this factor did, in fact, represent the development of quantitative skills.

The common factor in the halo effect model was responsible for more than half of the explained variance in the verbal skills items and two out of three of the quantitative skills measures. More than 90% of the explained variance in the personal development items was attributable to the common factor. This last finding suggested that the common factor was primarily representative of personal development, not satisfaction as was originally assumed.

Correlations derived from the multitrait-multimethod analyses also raised questions concerning the superiority of the halo effect model. These correlations are presented in Table 6 and all are statistically significant ($p < .001$). On average, there was no difference in the magnitude of the correlations for the two competing models. In addition, the correlations between the latent variables representing senior and alumni personal development exceeded 1.00 for both models. This finding, coupled with the relatively poor goodness-of-fit statistics for the alumni data, suggested that alumni reports of personal development were over-explained by both models.

Table 6:

Correlations Between the Latent Variables for Seniors and Alumni

	True				
	Satisfaction	Verbal	Quantitative	Personal	Humanities
Satisfaction	0.72				
Verbal		0.74			
Quantitative			0.74		
Personal				1.00*	
Humanities					0.82
	Halo				
	Satisfaction	Verbal	Quantitative	Personal	Humanities
Satisfaction	0.74				
Verbal		0.73			
Quantitative			0.73		
Personal				1.00*	
Humanities					0.82

* Correlation exceeds 1.00

Discussion

Before considering the implications of these results, it is important to underscore the limitations of this study. First, this research represents a single-institution, single-instrument study. Although the subjects in this study were representative of students at UTK, care should be taken in generalizing the results to other institutions, even other public research universities. Care also should be taken in generalizing these results to other self-report measures. Determining whether the observed positive

correlations between satisfaction and dimensions of learning and development in Pace's CSEQ or the scales developed by Pascarella and Terenzini are the product of a halo effect was beyond the scope of the present investigation.

A second limitation of this research was the use of a correlation matrix in the analyses. Using a correlation matrix instead of the more traditional variance-covariance matrix produces relatively weak evidence of measurement invariance. As Cudek (1989) has pointed out, invariance in correlation matrices does not necessarily imply invariance in variance-covariance matrices, since an infinite number of variance-covariance matrices can produce the same correlation matrices. As previously noted, reliance on correlation matrices as units of analysis in this study was necessitated by the fact that the measured variables were on an ordinal scale. Thus, there is a tradeoff in research on ordinal variables between utilizing the correct matrices for data analyses and the strength of claims that can be made about the invariance of the measurement model.

The third limitation of this research is the finding that the unique factors representing personal development for alumni were over-explained in both the true and halo effect models. This finding calls into question the appropriateness of the data, the models, or both. The fact that correlations exceeded 1.00 in the multitrait-multimethod analyses may be due to a multivariate outliers (Bentler, 1989), violations of multivariate normality of the inferred Y* variables (Wolfle, 1992), multivariate collinearity in the data (Kaplan, 1992), or model misspecification. Determining the exact cause of this problem was beyond the score of the present research.

In summary, it is impossible to unequivocally support either of the competing models. Although the goodness-of-fit indices revealed that the halo effect model provided the best fit to the data, the model did not produce a better explanation of the variance in observed measures. Nor did the halo effect model produce higher correlations between the same unique latent variables measured at two points in time. Despite the equivocal findings, the results of this study do indicate that the halo effect model is at least as good an explanation of the correlations among satisfaction and dimensions of learning and development as is assuming that the correlations represent a true relationship among the latent variables.

The findings of this study also called into question the assumption upon which the halo effect model was originally based. Initially, it was argued that the halo effect reflected a general sense of satisfaction with college. Results indicated that a relatively small amount of the variability in satisfaction was explained by the common factor representing the halo effect. Instead, the common factor explained the largest proportion of variability in perceptions of personal development. Interestingly, the role of the common factor in explaining personal development may actually support the halo effect. Symonds (1925) identified five conditions likely to produce large halo effects: (1) when the trait is not readily observable; (2) when the trait is not usually discussed in the classroom; (3) when the trait is not clearly defined; (4) when the trait involves relations with others; and (5) when the trait has high moral importance. Certainly the first four of these conditions are characteristics of perceptions of personal development. Thus, it may be more reasonable to expect that the halo effect would be greater for personal development than for satisfaction.

In addition to the findings regarding the plausibility of the halo effect, this research has significant implications for the assessment of goodness-of-fit in confirmatory factor analyses. In fact, the goodness-of-fit indices used in this paper have not, to this author's knowledge, been presented elsewhere. As traditionally formulated, the numerator in both the Relative Noncentrality Index (RNI) and the Tucker-Lewis Index (TLI) represent the decrease in noncentrality (improvement in goodness-of-fit) achieved by moving from a null model to a taget model. The denominator for the RNI is the null model, adjusted for sample size, and the denominator for the TLI is the reduction in noncentrality achieved by moving from the null model to a true model.

The indices used in this research differed from the traditional formulations in serveral important respects. Instead of using a null model based on unrelated measured variables, the present research relied on a null model representing uncorrelated latent variables (factors). This null model is similar to one proposed by Mulaik, James, Van Alstine, Bennett, Lind, and Stockwell (1989). Its use was based on the fact that the measurement structure for the learning and development items has been clearly established, and that the goal of this research was to investigate alternative models of relationships among latent variables.

The indices used in this paper differed from those suggested by Mulaik et al. (1989) in terms of what was considered the best (or true) representation of the data. Mulaik et al. recommended using a saturated path model as a measure of the best-fitting model. In this research, a saturated measurement model representing a perfect fit to the data was used. (This is the model recommended for the TLI.)

Thus, the goodness-of-fit indices used in the present investigation ignore changes in noncentrality associated with an accepted measurement model, focusing instead on changes in noncentrality due to different factor structures. The RNI represents the relative decrease in noncentrality achieved by a particular factor structure, while the TLI represents the ratio of the actual decrease in noncentrality to the maximum possible decrease in noncentrality for a given factor structure. The most significant characteristic of these indices is that they can be used to compare factor structures that are not nested.

Conclusion

It is worth reiterating that surveys play an important role in research on how college affects students and that questions dealing with students' perceptions of their growth and development are an increasingly important component in these surveys. While this paper does not unequivocally demonstrate that correlations among dimensions underlying satisfaction and reports of learning and development are due to the consistent error of the halo, it does show that the halo effect is a reasonable explanation of these results. Indeed, the halo effect is at least as likely an explanation as is assuming that the correlations among dimensions of learning and development are a true relationship.

Although the halo effect is frequently cited as a threat to the validity of performance appraisals, there is a surprising lack of discussion of the role of the halo effect in survey research. Texts dealing with survey research frequently address issues of how the wording of items may produce response patterns, but discussions of how the subject matter of the questions also produces highly correlated responses are lacking (see Bradburn, 1983). Even though the halo effect was originally proposed to answer questions about artificially high correlations among ratings made by an external judge, it is still a theoretically viable explanation for response consistencies in self-reports. After all, self-reports of learning and development are a form of self-rating.

The clear conclusion to emerge from this study is that researchers and assessment practitioners should exercise care in interpreting relationships among self-reports of learning and development during college, particularly as they relate to satisfaction with college. Observed correlations may well be artifacts of a constant error of the halo, rather than true relationships between satisfaction and dimensions of learning and development.

References

ACE/Winthrop national assessment survey results release. (1990, Winter). The Exchange, 1(5), 1, 13.

Aitken, N. D. (1982). College student performance, satisfaction, and retention: Specification and estimation of a structural model.

Alexander, J. M., & Stark, J. S. (1986). Focusing on student academic outcomes: A working paper. Ann Arbor: National Center for Research to Improve Postsecondary Teaching and Learning.

Astin, A. W. (1974). Measuring the outcomes of higher education. In H. R. Bowen (Ed.), Evaluating institutions for accountability (New Directions for Institutional Research Series, no. 1, pp. 23-46). San Francisco: Jossey-Bass.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. Research in Higher Education, 12, 155-187.

Bean, J. P., & Bradley, R. K. (1986). Untangling the satisfaction-performance relationship for college students. Journal of Higher Education, 57, 393-412.

Bentler, P. M. (1989). EQS: Structural equations program manual. Los Angeles: BMDP Statistical Software.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. Psychological Bulletin, 105, 467-477.

Bowen, H. R. (1977). Investment in learning: The individual and social value of American higher education. San Francisco: Jossey-Bass.

Bradburn, N. M. (1983). Response effects. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), Handbook of survey research. New York: Academic Press.

Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin, 105, 456-466.

Cudek, R. (1989). Analysis of correlation matrices using covariance structure models. Psychological Bulletin, 105, 317-327.

Cunningham, W. R. (1991). Issues in factorial invariance. In L. M. Collins & J. L. Horn (Eds.), Best methods for the analysis of change: Recent advances, unanswered questions, future directions. Washington, D.C.: American Psychological Association.

Ewell, P. T. (1984). The self-regarding institution: Information for excellence. Boulder, CO: National Center for Higher Education Management Systems.

Feldman, K. A., & Newcomb, T. M. (1969). The impact of college on students. San Francisco: Jossey-Bass.

Gold, K. F., & Muthén, B. O. (1991, April). Extension of covariance structure analysis: Hierarchical modeling of multidimensional achievement data. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Graham, S. W., & Cockriel, I. (1989). College outcome assessment factors: An empirical approach. College Student Journal, 23, 280-288.

Horn, J. L. (1991). Comments on issues in factorial invariance. In L. M. Collins & J. L. Horn (Eds.), Best methods for the analysis of change: Recent advances, unanswered questions, future directions. Washington, D.C.: American Psychological Association.

Jöreskog, K. G., & Sörbom, D. (1989). LISREL 7: A guide to the program and applications. Chicago: SPSS.

Kaplan, D. (1992, April). Collinearity diagnostics for covariance structure models. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Lenning, O. T. (1977). Previous attempts to structure educational outcomes and outcome-related concepts: A compilation and review of the literature. Boulder, CO: National Center for Higher Education Management Systems.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis. Psychological Bulletin, 103, 391-410.

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. Psychological Bulletin, 107, 247-255.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3d ed., pp. 13-103). New York: Macmillan.

Munro, B. H. (1981). Dropouts from higher education: Path analysis of a national sample. American Educational Research Journal, 18, 133-141.

Pace, C. R. (1987). CSEQ: Test manual and norms. Los Angeles: Center for the Study of Evaluation.

Pace, C. R. (1988). Measuring the quality of college student experiences. Los Angeles: Center for the Study of Evaluation.

Pascarella, E. T., & Chapman, D. W. (1983). A multiinstitutional, path analytic validation of Tinto's model of college withdrawal. American Educational Research Journal, 20, 87-102.

Pascarella, E. T., Smart, J. C., & Ethington, C. A. (1986). Long-term persistence of two-year college students. Research in Higher Education, 24, 47-71.

Pascarella, E. T., & Terenzini, P. T. (1978). Student-faculty informal relationships and freshman year educational outcomes. Journal of Educational Research, 71, 183-189.

Pascarella, E. T., & Terenzini, P. T. (1991). How college affects students: Findings and insights from twenty years of research. San Francisco: Jossey-Bass.

Pike, G. R. (1989). The performance-satisfaction relationship revisited: Specification and testing a structural model. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Pike, G. R. (1990, April). Dimensions of alumni perceptions of cognitive and affective growth during college. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Pike, G. R. (1991a, October). Dimensions of academic growth and development during college: Using alumni reports to evaluate education programs. Paper presented at the annual meeting of the Association for the Study of Higner Education, Boston.

Pike, G. R. (1991b). The effects of background, coursework, and involvement on students' grades and satisfaction. Research in Higher Education, 32, 15-30.

Pike, G. R. (1992, April). Students' perceptions of their learning and development at graduation and two years later: A preliminary investigation. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Pike, G. R. (in press). The components of construct validity: A comparison of two measures of general education outcomes. Journal of General Education.

Symonds, P. M. (1925). Notes on rating. Journal of Applied Psychology, 9, 456-461.

Terenzini, P. T., & Pascarella, E. T. (1977). Voluntary freshman attrition and patterns of social and academic integration in a university: A test of a conceptual model. Research in Higher Education, 6, 25-43.

Terenzini, P. T., Pascarella, E. T., & Lorang, W. G. (1982). An assessment of the academic and social influences on freshman year educational outcomes. Review of Higher Education, 5, 86-109.

Terenzini, P. T., Theophilides, C., & Lorang, W. G. (1984a). Influences on students' perceptions of their academic skill development during college. Journal of Higher Education, 55, 621-636.

Terenzini, P. T., Theophilides, C., & Lorang, W. G. (1984b). Influences on students' perceptions of their personal development during the first three years of college. Research in Higher Education, 21, 178-194.

Terenzini, P. T., & Wright, T. M. (1987a). Influences on students' academic growth during four years of college. Research in Higher Education, 26, 161-179.

Terenzini, P. T., & Wright, T. M. (1987b). Students personal growth during the first two years of college. Review of Higher Education, 10, 259-271.

Thorndike, E. L. (1920). A constant error in psychological ratings. Journal of Applied Psychology, 4, 25-29.

Valiga, M. J. (1982). Structuring the perceived outcomes of higher education. Paper presented at the annual forum of the Association for Institutional Research, Minneapolis.

Wolfle, L. M. (1992, April). Effects of postsecondary education on self-esteem: A lesson in LISREL. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.